RECOMMENDED PRACTICE

Analysis, Selection, and Implementation of Electronic Document Management Systems (EDMS)

An AIIM Recommended Practice Report prepared by the Association for Information and Image Management International

Approved
July 11, 2007

AIIM ARP1-2007

aiim
The Enterprise Content Management Association

AIIM ARP1-2007

AIIM Recommended Practice —

Analysis, Selection, and Implementation of Electronic Document Management Systems (EDMS)

An AIIM Recommended Practice Report

July 11, 2007

Abstract
This industry recommended practice presents a set of procedures and activities that should be considered and/or performed during all aspects of analyzing, selecting, and implementing electronic document management systems. This document provides a categorization of relevant national and international standards and reports, enabling users and organizations to quickly identify and locate required information for all aspects of the EDMS project.

Table of Contents

Tables

Figures

Foreword

At the time this AIIM Recommended Practice was approved, the Standards Board of AIIM had the following members:

Name of Representative	Organization
Linda Koontz, *Chair*	U.S. General Accountability Office
Leslie Banach	IBM Corporation
Robert Blatt	EID Inc.
Robert Breslawski	U.S. TAG to ISO TC 171
Betsy Fanning	AIIM
Bruce Holroyd	Eastman Kodak Company
Stephen Levenson	U.S. District Courts
Bill Manago	Computer Associates – MDY
Cheryl McKinnon	OpenText
J. Timothy Sprehe	Sprehe Information Management
Melanie Warfel	Adobe Systems Inc.

This standard was approved by Implementation Guidelines Committee. At the time this standard was approved, the Implementation Guidelines Committee had the following members:

Name	Organization
Robert Blatt	EID, *Committee Chair*
Richard Allen	New Jersey Dept of State
John Breeden	Virginia Dept of Transportation
Robert Breslawski	Eastman Kodak Company
Stacy Brown	Vistronix, Inc.
Terese Butler	California Department of General Services
Tony DeGruy	City of Stockton, California
Clark Emch	California Water Resources Control Board
Gustavo Espaillat	PKM Soluciones Digitales, S.A.
Theresa Finger	California Secretary of State
Jo-Ann Georges	University of the West Indies
Debbie Gearhart	Michigan Management & Budget Records Management
Mark Giguere	NARA
Arthur Hedge	CastleVentures
Robert Horton	Minnesota Historical Society
Kaz Jaszczak	Parascript
Jeff Kintop	Nevada State Archives
Derrick Lau	CCH Canadian Limited
Rick Laxman	LDS Church
Glen McAninch	Kentucky State Archives
Andrew Taylor	Georgia Dept. of Archives & History

Introduction

This document provides detailed information associated with the analysis, selection, and implementation procedures associated with Electronic Document Management Systems (EDMS). The development of this document is a result of organizational requests related to receiving industry, vendor-neutral information associated with industry standardized recommendations associated with standards, technical reports, guidelines, and best practices related to project activities.

For purposes of discussion, the terms document management and content management can be considered to be synonymous. As the Electronic Content Management industry (previously referred to as the Document Management Industry) has matured over the years, the ability to store electronic information has greatly expanded from only scanned images in the early 1980s to any electronic format managed by computers. As these capabilities became standard for EDMS products, the industry began changing the term Document Management to Content Management. When reviewing industry related information (e.g., technical reports, standards, marketing materials, etc.), readers should take into account that these terms are commonly used throughout the industry to refer to similar technologies and capabilities.

The first portion of this document provides detailed information describing each of these technologies, how they operate, and inter-operate. The second portion of this document provides detailed information associated with currently available industry standards and technical reports. The third portion of this document provides detailed information related to industry best practices associated with all the customary project phases for EDMS technology analysis, selection, and implementation.

As this information is focused on EDMS technologies and associated activities, other aspects of information management such as Records Management are not discussed in this document. It should be noted and acknowledged that a complete records management program is critical to any organization. For more information on other disciplines and records management technologies integrated into core EDMS foundational solutions, please refer to other documents available though the appropriate industry standard setting bodies.

Analysis, Selection, and Implementation Guidelines Associated with Electronic Document Management Systems (EDMS)

1 Scope

The scope of this AIIM Recommended Practice is to present a set of procedures and activities, which should be considered and/or performed during the analysis, selection, and implementation project phases associated with Electronic Document Management Systems technologies. This document will provide user level information outlining specific recommended activities to be completed throughout the various project phases typically performed when implementing these technologies. These steps and activities, along with compliance with relevant industry standards and guidelines should be examined and “certified” to ensure relevant technologies have been analyzed, designed, implemented, and managed, ensuring document/record validity when used in a business or government environment.

The term electronic document management used throughout this document is intended as an "all-encompassing" term referring to inputting technologies (scanning, indexing, Optical Character Recognition (OCR), forms, digital creation, etc.), management technologies (document services, workflow, and other work management tools), and storage technologies (primarily optical/magnetic). Additionally, this document will provide information to users related to what technical reports, guidelines, and standards have been developed for technologies commonly available in document management systems.

This document is not intended to be an all-inclusive paper on electronic document or content management and does not attempt to influence any single technology or provide legal guidance or legal opinions. While there are storage technologies other than optical/magnetic currently available (e.g., microfilm, microfiche, and hybrid storage systems) that are not included in this report, those technologies should be reviewed if determined to be appropriate by the end user organization.

1.1 Purpose

The purpose of this document is to educate and raise awareness related to planning, implementation, and management of web-based document management systems. It is intended to be from a vendor-neutral perspective and includes input from various state and county agencies responsible for mandating statewide or countywide procedures. As many public and private organizations throughout the United States are already in the process of planning or implementing these technologies, an industry standard guideline incorporating methodologies, approaches, and considerations from a wide range of governmental agencies and private industry can benefit all users.

1.2 Objective

The objective of this document is to define the topics and raise issues for each topic defined for the collective target audience: MIS staff, RM staff, vendors, integrators, and users.

1.3 Audience

This document is intended for anyone responsible for or interested in planning and implementing electronic content or document management systems.

1.4 Exclusion

This document is not intended to be an all-inclusive paper on document management and does not attempt to influence any single technology or provide legal guidance or legal opinions. While there are storage technologies other than optical/magnetic currently available (e.g., microfilm, microfiche, and hybrid storage systems) that are not included in this report, those technologies should be reviewed if determined to be appropriate by the end user organization. Technical reports and guidelines associated with these technologies are available from AIIM.

2 Normative References

The following normative documents contain provisions which, through reference in this text, constitute provisions of this Recommended Practice. For dated references, subsequent amendments to, or revisions of, any of these publications do not apply. However, parties to agreements based on this Recommended Practice are encouraged to investigate the possibility of applying the most recent editions of the normative documents indicated below. For undated references, the latest edition of the normative document referred to applies. ISO and IEC maintain registers of currently valid International Standards.

ISO/TS 12033, *Electronic imaging – Guidance for selection of document image compression methods*

ISO/TR 12037, *Electronic imaging – Recommendations for the expungement of information recorded on write-once optical media*

ISO 12651, *Electronic imaging – Vocabulary*

ISO 12653-1, *Electronic imaging – Test target for the black-and-white scanning of office documents – Part 1: Characteristics*

ISO 12653-2, *Electronic imaging – Test target for the black-and-white scanning of office documents – Part 2: Method of use*

ISO/TR 12654, *Electronic imaging – Recommendations for the management of electronic recording systems for the recording of documents that may be required as evidence, on WORM optical disk*

ISO/TR 14105, *Electronic imaging – Human and organizational issues for successful Electronic Image Management (EIM) implementation*

ISO/TR 15801, *Electronic imaging – Information stored electronically – Recommendations for trustworthiness and reliability*

ISO 15489-1, *Information and documentation – Records management – Part 1: General*

ISO/IEC 27002, *Information technology – Security techniques – Code of practice for information security management*

3 Terms and Definitions

For the purposes of this Recommended Practice, the terms and definitions given in ISO 12651 and ISO 15489-1 apply.

4 Electronic Document Management Technologies

4.1 General

Even in today's world, many organizations still function almost entirely in a "paper-driven" environment. This environment is a direct result of the need to maintain information on all aspects of the organization and can be seen throughout many organizations. Whenever considering EDMS technologies, organizations should consider implementing the necessary foundational components and then add other functionality as required by the business units.

Electronic Document Management Systems (EDMS) has become an all-encompassing term, referring to the integration of various underlying technologies including:

- Document imaging (used to convert hardcopy documents into digital format)
- Document/Library services (used to manage digitally born documents) (Note: Most EDMS systems allows users to use this technology to also manage scanned documents if desired)
- Workflow (used to route, track, and otherwise manage electronic documents and work activities)
- Enterprise Report Management (ERM) (used to store electronically formatted reports)
- Forms Management (used to incorporate interactive forms and manage related forms data)
- Optical Character Recognition (OCR)/Intelligent Character Recognition (ICR) technologies

Electronic document management systems provide users with greater access to digital information from a common user interface through the utilization of industry standard Internet browser technology. One of the primary reasons users prefer this level of technology is the distributed functionality and extent of digital information availability that can be accessed almost immediately after implementation.

EDMS technologies can be viewed as a set of building blocks as noted below in Figure 1. The lowest level is the operating system. Database Services and Storage Device Drivers are installed onto the server as the second layer. The selection of the database to be used is typically at the discretion of the organization, but is often standardized with Open DataBase Connectivity (ODBC).

Over the past 25 years, the industry has seen a significant shift from developing custom technology solutions at the database level to configuring/implementing commercially available software. As the EDMS industry and associated technologies matured, end user organizations were able to shift from a "development" model to a "configuration" model for the base technological components. From that perspective, the next layer in the "building block" is considered to be the EDMS server application. Early in the development and maturity cycle of EDMS technologies, end user organizations were required to provide database administration and resources. During the late 1980s and early 1990s, the EDMS technologies had not matured to the level approaching that of Commercially Available Off-the-Shelf (COTS). While the EDMS technologies were maturing, end user organizations were required to maintain the database along with the application.

Current versions of enterprise EDMS solutions have shifted the database administrative functions back to the vendor with the end user organization responsible for daily application maintenance and periodic server maintenance. Most enterprise EDMS solution providers provide their customers with technical support including system installation, initial configuration, application updates/patches, etc. This major shift from requiring significant technical resources at the end user level to vendor supported solutions has resulted in an extensive amount of vertical market penetration. This has been achieved through the use of the standard technology components configured to address specific environments and business needs.

Taking this into consideration, the next level in the EDMS building blocks are application modules that provide specific functionality. All enterprise EDMS solutions have at least one of these building blocks and in most cases include both Document Imaging and Document Services. Another module commonly incorporated into the enterprise solution is Workflow functionality.

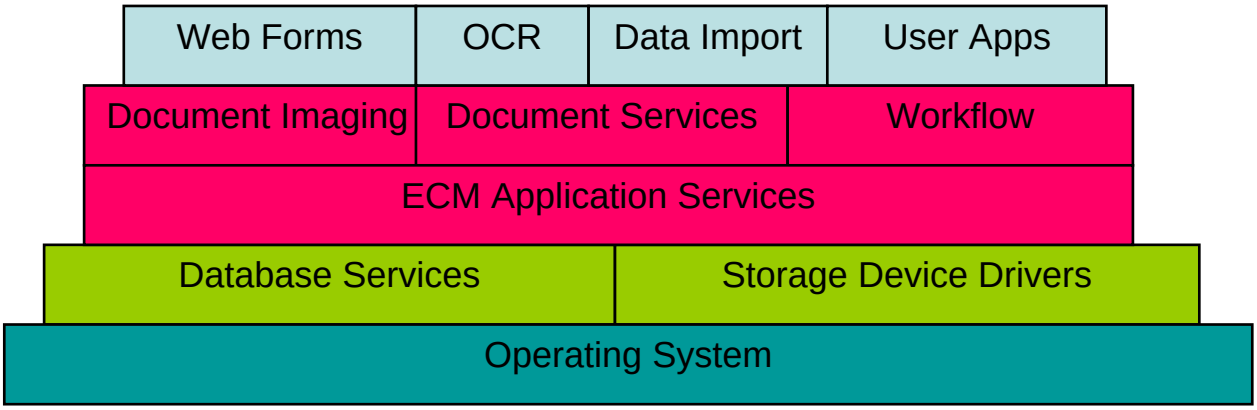

Figure 1: EDMS Technology Building Blocks

Workflow technologies can provide different levels of routing, tracking, and administration. These technologies can be grouped into three categories: administrative, ad hoc, and production. Administrative workflow is typically used by organizations where the processes do not change, or change very infrequently. Ad hoc workflow provides the ability for a user to create a “work process map” for a specific piece or type of work. Production workflow incorporates administrative workflow and ad hoc workflow capabilities along with providing extensive tracking and logging capabilities. When considering production workflow technologies, the organization should consider whether to use role-based or user-based technologies. User-based technologies require specific users to be assigned to specific tasks, while role-based technologies enable organizations to assign and re-assign users to groups or “roles” which are more easily managed.

4.2 Imaging Technologies

Document imaging technologies enable users to scan hardcopy documents into the system and store them in digital format. These technologies enable users to index or enter "metadata" into the system and they always utilize some form of storage technology to save the digital version of the document. There are four basic components to document imaging systems:

- input,
- identification,
- storage, and
- retrieval.

The **input** components typically consist of multiple single-sided (simplex) and/or double-sided (duplex) document scanners (or other input devices such as facsimile). The scanning stations are used to convert hardcopy documents into a digital format for subsequent storage and management in the document imaging system. The **identification** stations allow users to identify (or index) incoming documents allowing them to be retrieved at a later date. The **storage** part of the system consists of various components connected to the document management or workflow server and used to store, retrieve, and manage digital information. The **retrieval** part of the system consists of the user issuing a request for information that is then processed by the server. These requests are processed and the information is retrieved from the appropriate storage media connected to the server.

4.3 Document/Library Services Technologies

Document/Library Services technologies enable organizations to manage digitally born documents. Document/Library Services applications utilize applets, or thin-clients, to control the authoring, check-in/out, and/or version control of documents being developed, managed, or stored. This enables collaborative development when desired along with a mechanism to store/manage digitally born document libraries. The basic capabilities of these technologies include allowing authorized users to:

- load or import digitally born documents directly into the system,
- enter relevant metadata associated with the document,
- create virtual folders linking various documents together,
- check information in/out of the repository,
- make changes and check the modified information back into the repository,
- manage whether original documents are updated or replaced during the update operations, and
- establish security levels for groupings of documents.

The management portion of document/library services technologies includes the ability to restrict access to certain documents or groups of documents to only authorized users. Along with security controls, these technologies enable users to be granted different levels of access. For example, the author of a document might only grant read access to all users outside of a specific organization while granting "check-in/out" control to others who are working on updating the document. As the other users prepare to update the document, they would "check" the document out of the library, update the information, and then "check" the document back in.

Document/Library Services technologies ensure that any other user attempting to check the document out would firstly not be allowed to check it out, and secondly be notified that someone already has a copy being updated. Upon completion of the update cycle, the system automatically updates the version number of the document and makes it available to all authorized users.

Similar to Document Imaging, there are four basic processes associated with Document/Library Services technologies. (It should be noted that these terms may vary depending on various perspectives, e.g., records management vs. content management vs. archival management, but the underlying functions remain the same with slightly different terminology and/or descriptions.) The four are:

- import,
- identification,

— storage, and

— management/retrieval.

The **import** components typically consist of enabling users to import digitally born information into the system. This digitally born information can be any format/structure and can be loaded into the system in original or native format. Digital data does not need to be modified prior to being stored and should include all relevant indexing or metadata associated with the information. The **identification** components allow users to identify (or index) this digital information, allowing it to be retrieved at a later date along with providing a vehicle to store information related to the digital data itself (such as author, purpose, subject(s), and all types of information required by the end user organization to fully track all necessary metadata). The **storage** part of the system consists of various components connected to the EDMS or workflow server and used to store, retrieve, and manage digital information. The **management/retrieval** part of the system consists of the user issuing a request for information that is then processed by the server. These requests are processed and the information is retrieved from the appropriate storage media connected to the server.

4.4 Workflow Technologies

Workflow provides for the automation of business processes enabling users to control the process logic. This ability to control the various business processes enables mission-critical, content-centric business applications to operate in an environment otherwise cumbersome to implement and manage. This has resulted in most EDMS vendors offering an integrated workflow engine or integrating the workflow engine with various workflow products readily available throughout the industry. The primary difference between these two approaches is whether the product consists of only those components developed by the primary product supplier or whether the primary product supplier has integrated specialized technologies developed by other suppliers.

In the new approach to organizational networking, workflow is becoming a major tool in the automation of processes and posting of information to a web site. In these environments, workflow applications are becoming tightly integrated to legacy applications. The actual integration of workflow and other EDMS technologies has become more prevalent as various coalitions and standards committees and EDMS vendors have completed development of various standards.

The maturity of workflow technology and the associated trends are based on the separation of the processing rules from the processing scripts or work routing. In more sophisticated workflow environments, workflow scripts are tightly integrated to specific activities making the routing, editing, approval, and submissions of content manageable at the user level. Interaction with the various thin-clients would trigger sub-processes as defined in the workflow script, resulting in the appropriate applet being downloaded and/or launched.

Workflow computing is the automation of work processes performed throughout an organization. A workflow application automates the sequence of actions, activities, or tasks used to run the process. This includes tracking the status of each occurrence of the process and providing tools to manage the process. There are four basic components to a workflow system: processes, work queues, tools, and object data.

— **Processes:** An automated workflow application is made up of the different tasks or activities that must be completed to achieve a business goal. The workflow engine manages these processes. The workflow application works in conjunction with the engine to manage the work process.

— **Work Queues:** Work items are created and distributed according to preset rules and placed into work queues. Users or groups of users are assigned to various work queues as required for processing. Work items within these queues can also be automated.

— **Tools:** There are various tools accessed by the user including forms display, word processors, terminal emulators, legacy applications, etc. These tools are used to access existing host applications and perform office related activities as required to complete work.

— **Object Data:** "Object Data" is another term for any digital content referenced and/or used by the workflow system. The term became more prevalent after computing technology became sufficiently sophisticated to support video, audio, and other forms of information into the workflow system. These objects become the work item to be processed during the normal course of business.

4.5 ERM Technologies

Enterprise Report Management (ERM), which was previously known as Computer Output to Laser Disk or COLD, is an integrated software and hardware solution that stores and indexes formatted computer output (pages) on optical disk, magnetic disk, or magnetic tape as an alternative to paper printouts or computer output microfilm (COM). This formatted output consists of point-in-time reports, such as transaction listings of statements and invoices. Once this page output is stored on the ERM subsystem, it can be electronically retrieved, viewed, printed, faxed, and distributed to workstations and host computer terminals within organizations or throughout an enterprise.

While there are many different data types in the computing environment, the type of data which ERM technology is concerned with is typically the result of transactions (data files and database records) being formatted by the application into page-oriented form for printing on paper or computer output microfilm (COM). The structure and format of this output is known. This data is time-period focused—it is a snapshot of an internal system at a given point in time. These reports are often the basis for analysis or comparative reporting and they include the printed record received by users such as a statement or invoice. ERM systems have been designed primarily to handle this formatted output.

Essentially, the ERM process involves two procedures: recording (indexing and storing the data) and retrieving (making the data available to users). Within these two simple procedures, however, lay a myriad of complex tasks. Data must be downloaded or transferred to the ERM system server before it can be processed. The method used to transfer the data from the mainframe/host system to the ERM subsystem will vary depending on the communications capabilities currently in place.

Recording consists of writing new documents to the storage media and then making them available for retrieval. Recording speeds vary from system to system and are most critical in high volume systems. The recording process involves:

— transferring the data to the storage subsystem from the host,

— processing the pages from the transferred file (i.e., extracting index keys, compressing, and writing to optical storage), and

— adding the index records to the associated ERM database.

The retrieval process consists of the users accessing the system and selecting the appropriate report, or part of the report, for viewing. The selection of the information to be retrieved is based on information entered, by the user, into the query screen part of the viewing software. After the user selects the report, or part of the report to be viewed, the system retrieves the information, displaying it on the user's workstation.

4.6 Forms Management

The creation and utilization of electronic forms enable organizations to collect data in a standardized format and automatically enter or load the data into an EDMS solution. Electronic forms are typically created using either a forms design package or through the use of standard HTML editors. Forms design packages typically include not only the forms design components, but also enable organizations to "tag" or identify each field on the form and relate that data to a database or application that would receive and further process the information. These forms management tools also enable organizations to validate and/or perform edit checks on the forms as they are being completed to simplify data entry.

The usage of forms within the EDM industry has become widespread and most EDMS solutions incorporate some level of forms design and/or management as a portion of the standard product offering. In many cases, the use of forms design and management tools is replacing the older style of programmed forms that was required in the 1990's. By using these tools, organizations are able to quickly develop and deploy forms-driven data entry across the Internet without significant development efforts.

4.7 Optical and Intelligent Character Recognition

Optical Character Recognition (OCR) and Intelligent Character Recognition (ICR) enable organizations to quickly capture information from hard copy documents that need to be processed after document imaging and storage. OCR and ICR can greatly reduce the time required to index documents while enabling organizations to develop in-depth full-text searchable databases.

The value of using these technologies is especially evident when organizations need to capture specific portions of documents that are consistent. To capture this information, the utilization of "zoning" allows specific portions of similar documents to be identified and information within that "zone" captured and further processed as required by the application. Throughout the EDMS industry it has been found that the use of these technologies can also greatly improve the quality of information being indexed, while reducing the overall staffing requirements to perform the same functions manually.

5 EDMS Guidelines and Standards

5.1 Introduction

Industry guidelines and standards enable organizations to follow industry accepted practices and procedures. Standards and recommended practices specified in a federal, state, or local law or regulation are required specifically in the area covered by the law or regulation. Users wishing to require adherence to a standard or recommended practice should specify them in their procurement documents and contracts since this is the only way a vendor is required to meet a standard. Users of standards should also be careful to specify exactly what requirements in a standard must be met. It is possible for a system to “meet” a standard and still not deliver the required results if the contract is not specific about the contents of the standard or recommended practice.

Following industry guidelines and standards will further improve the ability of an organization to implement the selected technologies following policies and procedures found necessary, throughout the industry, to implement highly successful systems. These guidelines and standards also enable the organizations to implement products and technologies meeting their specific needs while being able to share information with other organizations who may, or may not, have the same product installed.

Industry guidelines provide specific information to users that will enable them to gain detailed information necessary to successfully prepare for, select, and implement the desired technology. The guidelines that users should evaluate include:

- Request for Proposal (RFP) guidelines;
- recommended document preparation procedures for scanning/indexing;
- planning considerations for technology implementation;
- how to determine what information should be included during document indexing;
- legal considerations;
- forms design;
- selecting the appropriate image compression methodology to be used;
- sampling procedures to verify information being scanned and indexed; and
- establishing quality requirements and quality control.

The relevant industry standards include those related to document services integration and toolkits, workflow integration and toolkits, document imaging related standards, and optical storage standards. Product suppliers must certify that their products meet the specified standard(s) to ensure that the product is, in fact, compliant with the relevant standard(s). It is important to note that as the industry creates and approves new standards and guidelines, this document will be updated to reflect those changes after the standards/guidelines have completed national or international approval processes.

5.2 Selecting the Appropriate Guideline or Standard

It is recommended that organizations preparing to select document management and workflow products review relevant industry guidelines and determine whether the vendors being evaluated meet the appropriate standards associated with that part of the technology. Titles of relevant standards and guidelines are provided in Annex A. Examples of various guidelines and standards for each of the technologies are documented below.

5.3 General

5.3.1 Industry Guidelines

Industry guidelines should be reviewed and will assist the organization during the preparation, planning, and implementation phases of the document management project.

5.3.1.1 Terminology

To gain an understanding of various terms used throughout the industry, the organization should review ISO 12651, *Electronic imaging – Vocabulary*. This document provides a detailed list of various terms that will be encountered during discussions with product vendors and integrators.

5.3.1.2 Human and Organizational Issues

When implementing these technologies, the organization will face several human and organizational issues. *Electronic imaging – Human and organizational issues for successful implementations of Electronic Image Management (EIM) implementation* (ISO 14105) provides detailed information gathered

throughout the industry related to system usability and adoption by the users. These guidelines will assist the organization during all the change management activities required for successful system implementation.

5.3.1.3 Request for Proposal (RFP) Development

Prior to selecting a specific product/integrator, the organization should document system requirements and provide them to those vendor(s), or integrator(s) being considered. Regardless of whether the RFP is being sent to a single vendor/integrator or multiple vendors/integrators, this document should be developed to enable the organization to clearly define their requirements and enable the vendor/integrator to clearly understand all business and technical goals and requirements.

5.3.2 Legality Issues

Legality issues that should be considered by legal advisors include information expungement, legal acceptance of records, retention requirements, and information redaction. Each of these issues is organization dependent and should be considered by legal advisors.

5.3.2.1 Expungement

It should be noted that information being expunged needs to follow specific legal rules and does not necessarily require that documents be permanently deleted, but can require that access to documents be permanently removed. Advice from legal counsel should be requested to determine whether permanent removal from accessing documents would meet expungement requirements. ISO 12037, *Electronic imaging – Recommendations for the expungement of information recorded on write-once optical media,* provides specific guidance.

5.3.2.2 Legal Acceptance of Records

Evidentiary issues associated with using electronic imaging systems and optical storage technologies need to be considered based upon local legal guidelines. Refer to ISO 12654, *Electronic imaging – Recommendations for the management of electronic recording systems for the recording of documents that may be required as evidence, on WORM optical disk,* for standard guidelines.

5.3.2.3 Retention Requirements

Users and systems designers should consult the organization's established retention requirement set forth in their Records Management policies and procedures. The system being implemented should ensure that it is able to retrieve the information throughout the required document life cycle. The storage media and its life expectancy rating must be considered, hardware and software obsolescence issues must be evaluated, and a sound migration strategy must be developed to ensure access.

Organizations that do not have current retention requirements should consider developing these documents. These documents enable organizations to manage existing records and provide a mechanism to automate when documents are to be archived, for how long, what action to take after the retention period is passed, along with numerous other organizational advantages from a management perspective.

5.3.2.4 Redaction

The process of redaction[1)] is elaborate, expensive, and subject to judicial review. It usually involves a careful, word-by-word examination of a document, the identification of the pieces to be "removed," the necessity of showing the location of the removed pieces, the inability of the document viewer to discover the redacted content, and the supervisory review and approval of the redaction—all with the recordkeeping to prove that the redaction was appropriate and conducted according to proper procedure. Therefore, the redaction process is usually done in a highly-controlled local setting. Redaction process software could be external to the document management system.

5.3.3 Technology Standards

Technology standards are developed for specific technologies and not at the "general" level. All industry related and relevant standards are listed in the appropriate technology category within this document.

5.3.4 Implementation Considerations

Implementation considerations should include system administration, security requirements, capacity planning, system performance, system scalability, and fax services. Each of these is discussed further in this section.

5.3.4.1 System Administration

When selecting the technologies required to support the business requirements, the organization should receive and maintain detailed information related to system administration functions required to administer and control all applications, security, system server hardware, and data backup/migration. The product supplier or system integrator should provide this information. These requirements should include:

- operating system management (updates, patches, backup, restore, etc.),
- application software (updates, patches, backup, restore, etc.),
- system security (user additions/deletions, security modifications, etc.),
- data migration (retention periods, media replacement, etc.),
- software trouble-shooting tools,
- hardware trouble-shooting tools, and
- database management utilities.

These technologies can be hosted off-site as well as on-site. The information technology (IT) group within the organization should, if desired, be provided the tools to perform these system administrative functions. At a minimum, the product supplier or system integrator should provide the ability for the IT group to manage the system and utilize the technical/support staff within the product supplier to resolve application and/or database issues that may be encountered along with assisting in software updates.

---

[1] "Redaction" refers to a process by which parts of a document are kept from disclosure. The parts might be such snippets as the name of a person, a Social Security Number, or entire paragraphs that reveal trade secrets. Documents might contain pieces of information that are protected by law from being revealed, e.g., because they contain privacy identifiers or trade secrets or other privileged information. In many redactions the rendition of the redacted document, whether hard or soft copy, will show a black bar through the space where the redacted content was located.

5.3.4.2 Security Requirements

To ensure the technology supports secure access that meets the organization's business needs, the solution must also be assessed with respect to how it supports end-to-end security as related to user authentication, document authentication, and secure network transactions over the Internet, Intranet, and Extranet as necessary. Understanding the complexity and scope of an organization's security issues especially when dealing with the Internet will require the collaboration of multiple organizational disciplines including legal, business operations, system administration, network administration, vendors, and external users of the system. For more information on security related requirements, organizations should review ISO 27002, *Information technology – Security techniques – Code of practice for information security management*.

5.3.4.3 Capacity Planning

To ensure an accurate assessment of the scope and size of the document management system, the organization should attempt to determine the capacity requirements of the expected solution. While the natural tendency of any organization is to regard all documents as essential and to store all documents with equal access, the selective categorization of documents by type, retention period, and frequency of access may contribute greatly to the final cost of the solution. Categorization of documents based on an organization's usage statistics is vital to determining how the document is finally stored in the system including optical library units and online storage. The following is a partial list of some sizing parameters to consider:

- system availability requirements,
- number of form types and documents,
- retention requirements by document or form type,
- frequency of document access,
- peak daily volume of new documents processed,
- volume of new cases for workflow consideration,
- number of internal users (case workers, researchers, data entry operators),
- number of users (local and remote),
- number of organizations, and
- number of remote sites.

5.3.4.4 System Performance

To ensure that information is available for use by the users within anticipated time frames, it is important that specific requirements related to performance expectations be defined. The organization should determine the anticipated response times they expect from the system for:

- document retrieval from long-term storage media,
- document retrieval from online cache,
- document viewing (over the LAN, WAN, Internet),

— document printing, and

— scanning/indexing performance.

5.3.4.5 System Scalability

Organizations should ensure that the solution be scalable. This scalability includes the ability to increase the number of processors in a multi-processor environment, increase the number of servers to operate in a cooperative fashion, as well as increase the storage capabilities as required by the organization. Requirements should include:

— the ability to increase the number of system users without component replacement,

— the ability to support other technologies, e.g., OCR, form management, etc.,

— the ability to support multiple servers and standardized non-alterable write-once storage solutions in a distributed manner, and

— the ability to support symmetrical multi-processing, if required by the organization.

5.3.4.6 Fax Services

The utilization of facsimile (fax) transmission services enables users to send and receive faxed documents directly at their workstation. When considering these services, organizations should evaluate the following requirements to support their specific business needs:

— **Outgoing fax without document viewing:** This provides users with the ability to fax documents directly from their computer without viewing each document first. The user should have the ability to select a range of documents and have them routed to the fax "server" for transmission.

— **Outgoing fax after document viewing:** This provides a user with the ability to fax a document during viewing. The user should have the ability to attach other documents to the outgoing fax as appropriate.

— **Incoming fax processing:** As incoming documents are received, the system should support the ability to receive incoming documents and automatically route the document based on configurable rules (via a system administration interface) either by incoming telephone number or through forms or OCR processing.

— **Fax status reporting:** The system should provide a fax reporting capability enabling users to view status and historical information related to faxes sent by the user. This historical reporting should be based on user security rights, preventing users from accessing other users' history, while supporting users with higher levels of security to access all historical records.

5.4 Document Imaging

5.4.1 User Guidelines

User guidelines should be reviewed for document imaging technologies to assist the organization during all project phases from the planning phases through the actual implementation. These guidelines will greatly simplify the overall process and ensure that critical aspects of preparation and implementation

planning are addressed early in the project rather than forcing the project to stall until planning issues are addressed.

5.4.1.1 Planning

During the planning stages of the project, the organization will need to address various issues including planning for the implementation of document imaging and preparing the documents to be scanned.

5.4.1.2 Indexing the Information

As the organization plans to implement document imaging, it should consider establishing relevant indexing field guidelines and procedures. These organizational guidelines should provide detailed information that should be considered when planning the indexing requirements for all current and anticipated documents to be scanned. Establishing all the necessary index values with the ability to add additional and/or other document types prior to system implementation greatly improves the value and quality of information being scanned and stored in the system.

5.4.1.3 Storage Technologies

During the planning stages of the project, the organization should review guidelines for the planning, implementation, and operation of long-term and permanent public records storage on electronic media, such as those in ISO 15801, *Electronic imaging – Information stored electronically – Recommendations for trustworthiness and reliability*. There are various approaches to document or record storage technologies including using optical WORM for long-term preservation to magnetic WORM for those organizations who need faster retrieval speeds and are less concerned about long-term archival issues. Another approach is to use magnetic WORM for temporary storage cache (where users can retrieve documents quickly) and optical storage for long term storage. Both sets of technologies are viable for document/record storage depending on user needs, regulatory storage requirements, etc. Users should exercise caution when using non-standardized or proprietary storage technologies.

5.4.1.4 Image Formats

The organization should ensure that all information being scanned or electronically received is stored in an industry accepted format such as TIFF, JPEG, JBIG, JPEG 2000, or PDF/A. Non-standard or proprietary file formats should not be used. Non-standard or proprietary formats include any formats used by a single vendor/source and accepted as a standard file format at either a national or international standards level. Proprietary file formats include "file-wrappers" used to encapsulate standard file formats within a non-standard structure.

5.4.1.5 Indexing Quality Control

As the system is moved into production, it will become important for the organization to develop a methodology of reviewing both index data and the actual documents to ensure the information is available and readable. The organization should establish a documented process to ensure that all documents are properly scanned and indexed. This documentation should be followed by all personnel performing scanning and indexing, along with providing a mechanism for index data entry verification prior to document committal to the storage media and/or transmission to the business process.

5.4.1.6 Scanning Quality Control

Scanning quality control measures enable users/operators to ensure that the scanner is operating within anticipated tolerances. ISO 12653 (parts 1 and 2), *Electronic imaging – Test target for black-and-white scanning of office documents*, provides additional information for production document scanners. Following these procedures will enable the user/operator to ascertain that the scanner is properly set up before scanning actual documents.

5.4.2 Technology Standards

Technology standards in this area are currently being developed. As these standards are completed and approved, those applicable will be incorporated.

5.4.3 Implementation Considerations

Implementation considerations should include document scanning (and indexing); scanning/indexing throughput; document image compression; post-scanning processing; document indexing; Optical Mark Reader (OMR), OCR, barcode, and ICR processing; quality control; query/retrieval display time; and printing times. Each of these is discussed further in this section.

5.4.3.1 Document Scanning

The document scanning part of the system should provide the ability for the users to quickly digitize documents and route these documents to the person performing the indexing operation. Requirements associated with this part of the system should include:

— the ability to support both batch processing and single document scanning and indexing;

— the ability to support document re-scanning;

— the ability to support both simplex and duplex scanning;

— the capability of the scanner to scan at the resolution meeting the specific image quality requirements of the system, such as 200, 300, or 400 DPI; and

— the ability to set page breaks when batch scanning fixed and variable length documents.

5.4.3.2 Document Scanning and Indexing

When implementing document scanning and indexing technologies, the requirements should include detailed information related to all processing phases. If color documents are to be scanned so that the image captures the color, the scanner must be capable of doing so. Patch code and barcode hardware and software should be included if these techniques are to be used for the automation of data indexing. When using these technologies, the user should be aware that barcoding and OCR technologies typically minimize key stroking during the indexing phase but do not always negate the need for manual indexing. The level of information captured automatically will vary depending on the quality of the incoming document and the ability of the system to accurately recognize the required information.

The issue of performance is of critical importance and the organization should ensure that the selected solution provides the ability to scan and index documents within anticipated time frames. The various processes associated with document scanning and indexing includes:

— the time required to prepare the document for scanning;

— scanning the documents, ensuring all documents and all sides (for double sided documents) are captured;

— the time required to index and verify the documents;

— the time required to route the document to the end user for further processing (if workflow technologies are being used);

— the ability to preset common fields (for indexing purposes) when scanning in batch mode; and

— the ability to support auto-indexing of documents using barcodes, OCR, or ICR.

5.4.3.3 Scanning/Indexing Throughput

The system must be capable of scanning either single or double-sided documents and must have scanners capable of processing the daily work volume at the selected scanning site. This processing will include document preparation, scanning, and indexing. The system must also be capable of supporting low, medium, and/or high volume scanning capabilities depending on user requirements and selected scanner. The total number of scan stations and indexing stations must be determined by the organization to ensure that all work can be processed within anticipated time frames and that stations are available for use when needed.

5.4.3.4 Document Image Compression

Image compression/decompression should support ITU Group 4, JPEG, JPEG 2000, JBIG, or other output format standards with no proprietary alterations of the algorithms. The selected compression technology should not include extraneous information unsupported by relevant industry standards. Users should be aware that when using proprietary file compression formats, the patent holder may require royalties and/or other fees, which are usually based on the total number of pages converted into that specific compression format, to be paid on a periodic basis. These licensing/royalty issues do not occur with non-proprietary formats.

There are various compression methodologies that are available. ISO/TS 12033, *Electronic imaging – Guidance for selection of document image compression methods*, is a guideline that enables users to select the appropriate compression technology that the vendor/integrator must support for different types of data. The different types of data may include scanned documents, line art, photographs, etc.

5.4.3.5 Post-Scanning Processing

Post processing may be used to provide image "clean-up" after the scanning and prior to indexing and final storage. This software generally performs de-speckling, de-skewing, and other functions to improve the quality of the scanned image with limited operator intervention.

Use of image “clean-up” and other post-scanning processing should only be used to improve legibility. Caution should be exercised when using these tools, as any material modification to the image may affect the ability to authenticate the document in a legal proceeding.

5.4.3.6 Document Indexing

Detailed information related to all aspects of document indexing should be clearly defined. This should include the ability for users to index documents on workstations other than the scanning stations and support the ability to:

— index images either prior to storage or immediately after storage,

— add other indexing values, and

— batch index.

5.4.3.7 Optical Mark Reader (OMR), OCR, Barcode, and ICR Processing

The main objective of the available recognition technologies is to reduce the amount of manual data entry for the capture of both hand-printed and machine-printed information from digitized documents. Although this technology will never eliminate the need for manual data entry, their effective use on targeted documents has produced remarkable benefits often evaluated in reduction of manual keystrokes. The following is a brief list of evaluation criteria to consider when analyzing the use of automated data capture:

- Is it possible to identify documents with sufficient volume to justify automated data capture processing? This is typically applies to forms containing both structured and unstructured content, but with identifiable information to be extracted.

- Is it cost effective? Determine the amount of data to be captured and the cost to support a manual solution, and then compare it to an automated data capture solution.

- Is it possible to re-design the target forms for improved recognition? The use of checkboxes, patch codes, barcodes, dropout ink, and OCR fonts all provide considerable improvement in recognition accuracy rates.

- How will the documents be batched for scanning? Will scanning use mixed form sizes? Will scanning use mixed form types? Is it possible to introduce a batch header sheet to streamline the scanning process?

- What are the business rules that may be used for post-recognition processing to improve the accuracy of the information captured? For example, the capture of a unique personal identifier can be used to automatically verify the name and address information against the organization's existing database.

5.4.3.8 Quality Control

When defining quality control for document scanning and indexing, the organization should include the ability for the user to:

- check and validate the complete scanning and indexing process,

- facilitate the re-scanning of poor quality images,

- verify readability of each page of each document,

- verify proper indexing of each document,

- verify accurate page counts for each document, and

- verify accurate security for each document.

5.4.3.9 Query/Retrieval Display Time

Query and retrieval display time is commonly of high importance to the users. The users should define the anticipated performance requirements prior to system design and hardware procurement. These performance requirements should include maximum response times anticipated during production taking into account the total number of anticipated simultaneous user requests; the total number of drives; whether the information is available in an on-line, near-on-line, or off-line mode, etc.

These time periods include all time required to retrieve the appropriate optical/removable media (when necessary), reading all requested pages from storage media, storage of all requested pages on magnetic cache (if being used), and subsequent transmission of the first page to the user for viewing. When removable media (e.g., optical WORM, CD, DVD, tape, etc.) is implemented, this response time should take into account the time required to "spin" the drive down, eject the media, retrieve new media from the storage bays, insert the media into the drive, "spin" the drive up, and retrieve information from the media.

5.4.3.10 Printing Times

The imaging system must be capable of printing user selected documents within anticipated user established time frames. This response time includes document retrieval from optical storage and transmission to the selected printer. The user should have the ability to select a document, or range of documents, to be printed without being forced to view any of the pages prior to print submission.

5.5 Document/Library Services

5.5.1 Technology Standards

Technology standards should be evaluated by the organization to determine which standards are important and relevant to the overall project goals and objectives.

5.5.1.1 Open Source Distribution

The product vendor/supplier should certify that the organization can use open source document services software and metadata definitions (information describing the document) with the specified product. This will enable the organization to integrate other document services technologies without significant system re-development.

5.5.1.2 Development Toolkits

The product/vendor supplier should certify that the system uses industry standard application programming interfaces. This will enable the organization to implement a document services system and access information stored on other document services implemented throughout the network. These toolkits simplify application development and will enable the organizations to develop a common user interface regardless of the product used to “house” the actual data.

5.5.2 Implementation Considerations

Document services enable users to create, modify, and manage electronic files typically associated with various office processing applications. These capabilities include version control/check-in and check-out; logical folders; group/user security; document security; PDF/A, HTML, and XML conversion; and document publishing to a web site. Each of these is discussed further in this section.

5.5.2.1 Version Control/Check-in and Check-out

The organization should ensure that the product fully supports version control and check-in/out methodologies. Version control should automatically update the version number when a previously "checked-out" document is returned to the information repository. The system should prevent more than one person from checking documents out for modification and use a security model ensuring that only authorized personnel can perform these functions.

5.5.2.2 Logical Folders

The ability for the users to "logically" link a single document to multiple folders is important to prevent document duplication. The organization should ensure that the selected product supports the ability for an authorized user to create a copy of a document within a specific folder, or set of folder(s), while maintaining only one physical copy of the document within the system. The system should provide information related to which folders are "linked" through a query mechanism available to authorized users.

5.5.2.3 Group/User Security

The system should provide the ability for organizations to apply security access/restrictions at both the group and user levels. The group level security should apply to all users within the defined group, while user level security should provide additional security restrictions or capabilities for specified users.

5.5.2.4 Document Security

The system should provide the ability for organizations to apply security at the document or file level. Only those users with appropriate security levels should have access to these documents and/or files. This security should include read, update, annotation, highlighting, “mark-up,” and creation control.

5.5.2.5 PDF/A, HTML, and XML Conversion

The system should provide for PDF/A, HTML, or XML data conversion as required by the organization. This conversion should enable the users to convert existing office documents into a standardized format that can be accessed through a standard web browser.

5.5.2.6 Document Publishing to a Web Site

The system should provide the ability for an organization to update an existing web page automatically, after completion of a review/approval process or manual review and conversion by the “webmaster.” This document publishing functionality should include the ability to store native file formats or utilize web templates to reformat the document into either HTML, XML, or PDF/A format.

5.6 Workflow

5.6.1 Technology Standards

Technology standards have been developed by the Workflow Management Coalition (WfMC) into a Workflow Reference Model. The significant aspects of the Workflow Reference Model can be summarized into the following three categories, each building incrementally on the preceding:

1. A common vocabulary for describing the business process and various aspects of the supporting technologies to facilitate automation. This provides the essential foundation for the subsequent detailed discussion on how a workflow system could be architected in a general sense.

2. A functional description of the necessary key software components in a workflow management system and how they would interact. This was developed in a “technology neutral” manner to allow the model to be independent of any particular product architecture and implementation technology.

3. The definition, in functional (or abstract) terms, of the interface between various key software components that would facilitate exchange of information in a standardized way, thus enabling

interoperability between different products. Five such interfaces were identified and became the foundation for the WfMC standardization program.

An important principle was that the Reference Model focused specifically on workflow management technology and standards. It deliberately did not attempt to define standards in related areas in which other industry bodies were working; these were seen as complementary.

The Five Interfaces

Each interface was initially specified as a business level statement of objective, that is to say what the interface was intended to achieve in business terms and why a standardized approach was desirable. This was subsequently followed by a detailed, but abstract specification of how the interface operated and finally (for most interfaces) a "binding" specification covering the implementation of the interface in a particular technology.

Interface 1 was developed to support the exchange of process definition data between BPR (business process reengineering) tools, workflow systems, and process definition repositories, enabling users to select the most appropriate tool for different aspects of the business process lifecycle. It was specified as a Process Definition Meta-Model, defining the process objects, their attributes and relationships, and a textual grammar for expressing the process definition structure and information content. This was subsequently re-expressed as an XML document definition (XPDL).

Interface 2 was developed to facilitate client application integration with different workflow systems, in particular to support the principle of (client) application portability and reuse with different workflow management systems. It was specified as a series of Workflow APIs (WAPIs) to allow the control of process, activity, and worklist handling functions. These were originally defined in "C" and subsequently re-expressed in IDL (as part of the OMG workflow management facility) and OLE. A set of "C" APIs for manipulating process definition objects and attributes was also defined.

Interface 3 was scoped to provide a common framework for third parties to integrate other industry applications and services, including specific support of agent interfaces to provide a common framework for access to legacy applications. It was developed as a set of five basic API calls, defined within the WAPI document to support a common mechanism for connection, disconnection, and calling to a variety of agents or other third party software environments.

Interface 4 was developed to facilitate process automation across multiple heterogeneous implementation environments. It comprises an interchange protocol covering five basic operations, specified in abstract terms (initially defined in IDL) and with separate concrete bindings. The initial version was defined as a MIME body part for use with e-mail; subsequent versions have been specified in XML (Wf-XML). Ongoing work has led to version 2 of Wf-XML, layered over SOAP and ASAP.

Interface 5 enables consistent audit and administration of workflow cases across systems, through the specification of a common model for audit data, including event identification, formats, and recording. As such it was originally specified in abstract terms, although a set of common APIs for access to audit data was subsequently developed. Recent work is aimed at expressing the audit data structure as a set of XML structures.

Although conceived as five individual interfaces, the separation is apparent only when viewed in the context of the stated business objective. In reality there is significant commonality of function between the various "interfaces;" for example the triggering of the initiation of a process execution is fundamentally the same action whether it is done client side (i/f 2) or server side (i/f 4). The evolution of the WAPI (API) specification started with client application interactions but expanded to include a full repertoire of API calls. Similarly, Wf-XML was developed initially for server-server interaction but has also been used successfully for client-server interactions.

A more useful and fundamental distinction is perhaps to take a view of each interface from the perspective of process ownership and administration control. In particular, interfaces 2 and 3 may be considered to be “tightly bound” to the local workflow management system and reflect a local view of resource management—interface 2 handling interaction with human resource and interface 3 interaction with automata resource. This has two significant consequences. In the first place the process definition is localized to the point of process enactment through the expression of the resource assignments (e.g. participants and applications). Secondly the Reference Model could make the simplifying assumption that specification of messaging between a WFMS and participants need not be contained in detail within the process definition. It becomes a function of the WFMS locally to organize the most appropriate form of interaction with the participants via local *Worklists* (web access, e-mail, etc), according to the defined (within the process definition) *Activity* or *Procedure.*

5.6.1.1 Workflow Development Toolkits (WfMC Interface Specifications 2 & 3)

The vendor should certify that the product supports Workflow Application Programming Interfaces (WAPI). These APIs, as described in Workflow Management Coalition (WfMC) documents, ensure the implemented product provides a consistent method to access workflow management functions particularly in cross-product implementations.

5.6.1.2 Workflow Auditing (WfMC Interface Specification 5)

The vendor should certify that the product supports the WfMC audit specification. This specification details information to be captured and managed by the workflow system during operation. This will ensure that all relevant data is associated with all functions within the workflow technology.

5.6.1.3 Workflow Interoperability (WfMC Interface Specification 1)

The vendor should certify that the product supports industry interoperability standards including the usage of standard e-mail systems. These interoperability standards will enable the organization to share workflow information directly between different workflow systems without requiring specialized development.

5.6.2 Implementation considerations

Implementation considerations should include workflow, role versus user, routing requirements, graphic "rule designer," work monitoring, escalation procedures, error handling, and time-out procedures. Each of these is discussed further in this section.

5.6.2.1 Workflow

Workflow technologies include various types of routing including ad hoc routing, administrative routing, and production routing. Ad hoc routing enables the user to specify a specific process for a document to follow for that document only. Administrative routing enables users to define specific routing for a specific type of work that is always followed, regardless of the data within the work being routed. Production routing enables the users to define rules and work methods based on the document type and data contained within the work item. As the data changes, the production routing system should process the document accordingly, including the ability to support work timeouts, escalation, and work reassignment.

5.6.2.2 Role versus User

There are two approaches to defining users within a workflow environment. The first method is to define a specific user to manage a specific task or activity. The second approach is to define a role within the work task or activity and then assign as many users as necessary or appropriate. Organizations should require a "role" based system when implementing production workflow technologies.

5.6.2.3 Routing Requirements

For those organizations requiring production workflow, the system should allow a user to route a document to another user. The following capabilities should be considered:

- the ability to automatically route documents into a routing queue based on document type or "type of work,"
- the ability to support multiple routing queues for each user based on the "type of work,"
- the ability to sort/retrieve documents in a routing queue in date order,
- the ability to sort/retrieve sections in a routing queue in "type of work" order,
- the ability to sort/retrieve documents in a routing queue in document type order,
- the ability to sort/retrieve documents in a routing queue for a specific person,
- the ability to change a “pre-defined” routing path,
- the ability to "pend" or "hold" items in that user’s routing queue for work at a later time,
- the ability to retrieve specified documents from the routing queue on demand,
- the ability to define which documents require additional documents prior to forwarding,
- the ability to define timeframes for when additional documents must be received,
- the ability to define action to take if specified documents are not received by specified date, and
- the ability to process defined documents as a “logical” folder.

5.6.2.4 Graphical "Rule Designer"

The system should support the ability for authorized users to create and modify work rules associated with the workflow system. This ability should include graphical based design and management tools that would be used to create/modify work rules within a Windows or browser based user environment.

5.6.2.5 Work Monitoring

When selecting workflow technologies, the organization should evaluate whether work monitoring is required for their operation. Work monitoring tools enable the users to monitor current ongoing work, typically in a real-time basis. This work monitoring is used not only for "load-leveling" of ongoing work activities, but also to see if there are any "bottlenecks" in the overall workflow process.

5.6.2.6 Escalation Procedures

For those organizations requiring production level workflow, the selected solution should include the ability to automatically route work to a different user based on a specific rule or set of rules. The solution should also include the ability for users to manually escalate work as appropriate. During this escalation procedure, the solution should have the ability to have the work item returned or permanently reassigned as determined by the user.

5.6.2.7 Error Handling

As workflow items can include information not previously anticipated during the rules definition, the organization should require that the solution include the ability to handle errors within the routing of work through the workflow engine. The error handling should include the ability to pre-define a role that would receive the appropriate work items that are determined to be in error.

5.6.2.8 Time-out Procedures

When workflow is implemented, there are many instances where the timeliness of completing a specific work activity, or group of activities, is important. The ability to establish timers for all work items becomes very important. The organization should require that the solution support "timer" mechanisms and that the user is able to set these time-out values for specific activities throughout the graphical work "rule designer" tool.

5.7 COLD/ERM

Industry standards and user guidelines in this area are currently being developed. As these standards and guidelines are completed and approved, those applicable will be incorporated.

6 Best Practices Associated with ECM Project Phases/Activities

6.1 General

The steps and/or activities that provide ECM industry accepted guidance to assist end user organizations in all aspects of the ECM project include but should not necessarily be limited to: business/operational process review, business/operational/technical requirement documentation, evaluation and selection of appropriate technologies addressing specific business issues, conversion/data issues, records retention and management issues, and implementation activities. Each of the following sections provides detailed information on those activities requiring completion prior to product/vendor selection. There are numerous steps and procedures associated with analyzing business requirements for the identification and selection of relevant technologies to be considered for implementation. Annex B provides a high level check list of *Best Practices Associated with ECM Project Phases/Activities*. This annex provides a high level listing of recommended project steps/activities associated with a full EDMS project, beginning with process analysis through system/product implementation.

6.2 Process/Procedure Baselining

The purpose of process/procedure base lining is to clearly define existing processes/procedures and identify issues and problems currently encountered. This is achieved through a detailed analysis of existing processes and procedures. When performing this analysis it is important to capture and document activities including:

- how documents and information are received;
- what occurs to these documents after receipt (e.g., stamping, sorting, logging, delivery, etc.);
- how these documents are used and how many people use the same document to complete a specific activity or process;

- what happens to the document during the processing (annotation, highlighting, copying, etc.);
- after the processing is completed, where the document is stored, whether there are multiple copies, etc.; and
- how established document retention timeframes and the process of document destruction after reaching the destruction date within the retention policy are adhered to.

This information should be gathered through interviews with selected users within each processing unit. These users should include experienced users (non-management) and management personnel. It is important to note that the team gathering this information should represent the business units from a user perspective and include all processes and procedures currently being used. As the baselining process continues, users may describe processes and/or procedures that are not "officially sanctioned" in the day-to-day processing. These workaround, or alternative methods, need to be documented, as well as all other user workarounds and methodologies implemented to complete daily work activities.

Upon completion of this documentation, the users should have an opportunity to review the baseline document to ensure that all functions and activities related to their processing have been accurately captured and documented. It is very common for these documents to have multiple versions presented prior to user sign-off. This is due to the fact that most users do not have complete documentation at the detail level related to how the documents are managed.

There are three basic activities recommended throughout the ECM industry related to process baselining. These activities include: the development of a high-level baseline establishing the overall structure of the business process, the detailed baseline documenting specifics of each task/procedure identified during the high-level baselining, and associated manual processing metrics.

6.2.1 High Level Baseline

The first step in documenting business processes is to develop a high-level (management view) of the manual or document based business processes. Selected and representative users from organizational staff should be interviewed to identify all general work activities, policies, and business procedures. These processes should be documented in a “graphical” format developing the “high-level” process schematics documenting the manual or document based processing flow throughout the organizations. These “high-level” schematics are then further “exploded” during the detailed process documentation.

6.2.2 Detailed Process Baselining

Upon completion of the high-level processing schematics, those areas identified by managers and supervisors are further examined. Selected users are interviewed to obtain the detailed processing perspective. These interviews should include discussing how work and work-related information is received, processed, and “moved” between groups, departments, and other users.

These processing schematics should be documented using a graphical tool enabling the organization to review the processes in an interactive fashion on their computers. A read-only version of the software should be provided to any organizational resource with the need to review this documentation. Users identified both during the initial project planning stages and throughout the process should be interviewed to ensure that all relevant portions of the business processes are documented, showing all hand-offs and any other work-related activity.

During the process baselining activity, processing rules and conditions are identified as decision points and document routing/hand-offs as the document moves through each identified process. Rules and conditions are considered to be those decision points and hand-offs that dictate how information flows through the process. A small example showing the level of information gathered detail for each process including “rules and conditions” is shown below in Figure 2.

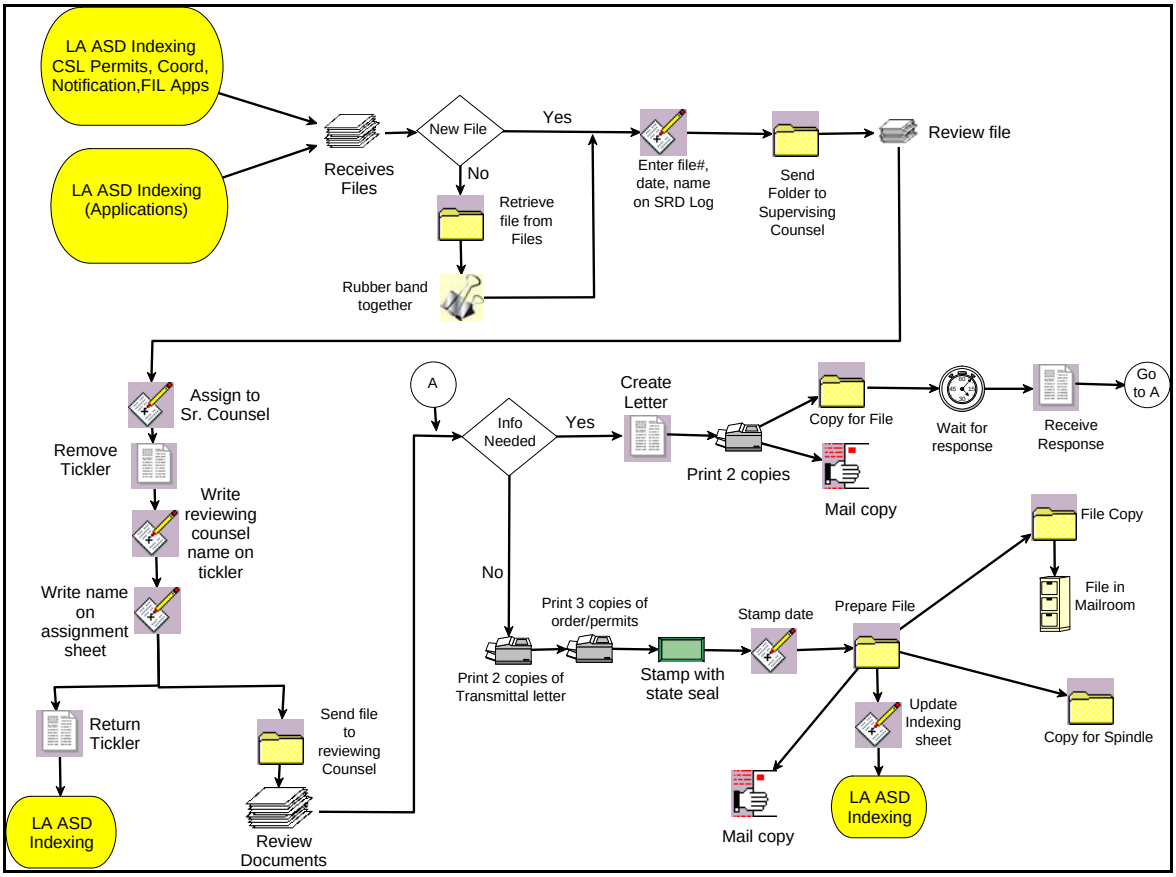

Figure 2: Sample Detailed Process Schematic

6.2.3 Processing Metrics

A Processing Metrics Report should be developed that would include information related to personnel time currently spent on all work related "manual" activities identified through the high-level and detailed schematic development. These tasks can then be further examined by organization management to identify those processes and/or activities that can be potentially replaced/enhanced with technology and those processes/activities that can be replaced/enhanced through organizational change management.

This enables the organization to evaluate how much time is spent managing the current workload along with anticipated time after new/updated technology implementation. Time associated with processes/activities identified in the schematics should be documented to identify items similar to:

- time spent logging receipt of documents;
- time spent copying, filing, and locating documents;
- time spent manually managing digital documents from creation through storage;
- time spent manually routing and tracking documents as they are processed;

- number of personnel performing associated major processes/activities;
- classification of documents and associated document volumes; and
- estimated number of multiple copies of identical documents throughout the organization.

As accurate time-related information is typically difficult for users to gather, it is common for an organization to follow a very conservative approach, evaluating how much "work" time is spent on these activities on a daily or weekly basis and then reviewing the information prior to inclusion in this documentation, while comparing the received information with overall times for the processes and other aspects of organizational work related activities.

6.3 Anticipated Processes/Procedures

Upon completion of the baselining process, this information is evaluated to determine where both non-technology based and technology based changes could be implemented.

Examples of non-technology-based change typically include reduction in document copies, the revision of outdated procedures, elimination of redundant procedures, and elimination of duplication of processes/procedures between organizations. During the non-technology change review, end user organizations should consider the impact on existing operations when updating/changing processes/procedures.

Examples of technology based change typically include automated logging of document receipt, automated routing for processing, and detailed history related to work activities associated with each work item or document. Additional examples of technology based change are related to which technologies are determined to be of benefit to the organization and how those technologies would be implemented, resulting in a different method of conducting business.

The level of detail associated with the new or anticipated processes should take into account how legacy systems would be updated, modified, and/or replaced. It is very typical to have a high number of personal databases, spreadsheets, and other tracking/data tracking activities that can be consolidated into the ECM system. It is also very common to have new processes and procedures established to not only manage digitally born information, but to support electronic distribution of this information to users within the organization and to users external to the organization through the use of electronic transmission.

6.4 Technology Requirements Definition

After identifying the relevant technology based changes required by the organization, solution requirements should be documented providing detailed information to potential solution vendors. This document should clearly define anticipated user and system functionality in sufficient detail enabling potential solution vendor to understand the business problem/issue being addressed and desired results after solution implementation.

When developing the solution requirements, organizations should consider documenting desired and/or required document management functionality and capabilities identified throughout this document.

6.5 Document Classification and Indexing Model

An important aspect of any ECM effort is to establish an enterprise wide document classification and indexing methodology. This effort should include meeting with all appropriate staff identified by the organization as being representative of how the ECM solution would be used to discuss existing document classifications and to prepare a full classification plan that can be implemented by the selected

solution integrator/software provider. It is important to note that when implementing an enterprise foundation the resultant document classification and indexing methodology should:

- be flexible and able to be updated as required by the organization without adversely affecting documents already stored by the solution;
- fully addresses records retention issues and schedules;
- enable staff to easily locate and retrieve documents;
- support the ability to implement document security;
- provide logical information groupings; and
- support the use of "virtual folders" where staff can search and organize documents during the retrieval process without forcing document replication or copying.

This classification should be fully compliant with the organization records retention schedules and link to the records management classification and structure. This will enable the business side of the organization to further manage the business documents and records as required by the records management team.

6.6 Business Objectives and Requirements

The business objectives, functional requirements, and expectations should be clearly defined. The business/functional requirement documentation should include both technology driven and non-technology driven requirements and detailed information related to:

- Business objectives of the project
- Business functional requirements
- Business expectations

This document should contain specific information related to current and near-future business needs and requirements identified through the business baseline activities, and interviews with business, technical, and management team members of the organization.

Additionally, the definition of Critical Success Factors (CSF) associated with how the resultant solution would be evaluated to determine overall achievement of these objectives and requirements should be clearly defined. CSFs are those items considered to be either a business or technical requirement for the organization. The CSFs should enable the organization and the vendor to identify those areas of critical importance related to the successful implementation of the desired technologies.

Common examples of critical success factors from both a business and technical perspective typically include:

Business Related Goals:

- **Improved service:** Users need the ability to quickly access and review information managed by the document imaging and workflow system.

- **Ability to track and monitor work activities:** The system should enable the users to track all ongoing work including the ability to re-assign work from one user to another. This tracking capability will enable the organization to implement workload leveling when appropriate.

- **Centralization of historical information between organizations:** The system should enable the organization to maintain centralized history related to all activities associated with the client/constituent. This history centralization should include both system-generated activities (e.g., date scanned, date routed, etc.) and user generated information such as notes taken during telephone conversations. The users should have access to information allowed by their security access, limiting access to information required by higher levels of security.

- **Increased efficiency of available resources:** The organization should be able to use the selected technologies to support ongoing business activities. The selected technology should enable users to decrease time spent on paper and file handling activities—including stamping, stapling, copying, delivering, and filing documents—and increase time in the areas of work processing.

- **Satisfying organizational and/or government regulations pertaining to document retention:** The use of electronic data storage must adhere to any laws and/or regulations covering the storage, retention, and retrieval of information on electronic storage media.

- **Decreased storage costs:** The solution must provide the ability to use optical storage technology to reduce the overall cost of storing and retrieving all “hardcopy” information.

- **Decreased costs for manual document management:** The cost for manual document management should be reduced along with an increase in the ability to provide improved service at a lower cost per request.

- **Simplified user access to application, work-order, and other data:** The overall solution must enable the users to quickly select and access the desired information without using highly complex user interfaces or tools. The user interface needs to be easy-to-use by the various system users.

Technical Goals:

- **Scalability:** The system must be fully scalable, allowing for an increase of the number of users and volumes of data without replacing primary system components. This scalability must be in the areas of increased memory, disk storage, optical storage, CPU speed and size, etc.

- **Migration path:** A clearly defined migration path must be fully supported by the proposed solution. This migration path must provide for the integration of new document management technologies to ensure proper integration without adversely affecting the proposed solution and/or data managed by the existing system(s).

- **Modularity:** The various client-based applications must be modular allowing for implementation of additional functionality without adversely affecting the overall system solution. This includes the ability to add routing; "virtual" file folders, high-volume printing, automated fax services, workload distribution, monitoring, etc.

- **Browser based access:** The system must fully support browser based technology while the various web servers will provide all the necessary mechanisms to store and retrieve information requested by the user, system level security for both users and data, and associated system management functions. All applications must be fully integrated to prevent redundant hardware and software on both the workstation and web server platforms.

— **Use industry standard components (no proprietary architectures allowed):** The associated components within the solution must be commonly available throughout the document imaging and workflow industries, be fully supported by the selected product supplier, and have full user and/or development documentation and libraries.

6.7 Technology Evaluation Guidelines

When evaluating appropriate technologies required or necessary to meet business and technical goals, the organization should consider several factors associated with the technology. The evaluation of the appropriate technology should include:

— **COLD/ERM:** When evaluating COLD/ERM technologies, the organization should review the downloading, indexing, and storage processing requirements. Additionally, the organization should consider the complexity of configuring the system to support new and/or modified report formats and indexing requirements. The ability of the technology to support simplified user access to data via a "query" screen and the ability to "cut and paste" information from a retrieved report or page to a standard office application should be considered. When evaluating COLD/ERM technologies, the organization should ensure that the system is capable of loading and indexing the daily work volume without impacting the users. This functionality of "loading" should include automated indexing based on templates defined by authorized users.

— **Document imaging:** When evaluating imaging technologies, it is helpful for the organization to perform "site visits" to other organizations similar in size and processing who have implemented the solution being considered. The purpose of these site visits is to gather information related to issues/problems encountered by other users that potentially have not been identified or addressed by the organization. During these site visits, all aspects of document scanning, indexing, and verification should be discussed. Overall system performance should be reviewed along with ease of use and processing accuracy and organizational satisfaction with the product/solution and the product/solution provider.

— **Document services:** These services enable users to manage electronic information independent of the tool used to create the information (e.g., word processing, spreadsheets, facsimile documents, etc.). Document services typically enable users to check documents "in" and "out" of information repositories; support document version control; and support document, group, and file level security rules. When evaluating these technologies, the organization should consider whether the product supports these functions along with their integration with web publishing components (described below).

— **Workflow:** When the organization determines that workflow technologies are required, it must be decided whether ad hoc, administrative, or production level technologies are required. For ad hoc and administrative routing/workflow requirements, the organization should evaluate whether the product includes simplified authoring tools (for non-complex routing procedures), which can be used in a graphical environment along with monitoring capabilities. The monitoring capabilities should enable authorized users access to work queues or "baskets." These administrative and monitoring tools should further enable the authorized user to re-route work items and establish basic escalation and "time-out" procedures. These escalation and 'time-out" procedures enable the users to establish a specific amount of time that a work item can remain at any specific activity, or establish a total amount of time to elapse prior to automatically sending the work to a specific person or role. When the organization determines that production level workflow technologies are required, the escalation and "time-out" requirements should be included, but additional functionality should be considered. This additional functionality should include the ability for authorized users to build complex workflow rules and support load-leveling functionality and real-time work queue or "basket" monitoring.

— **Automated Data Capture:** In many situations, the inclusion of OCR/ICR technologies can be justified solely on the reduction of manual data entry costs associated with indexing and capture of specific content from scanned documents. As there are many data capture products available that can be integrated with most document management systems, the organization should pay particular attention to the expected benefits and the ability to measure these benefits during the evaluation. When evaluating OCR/ICR/barcoding technologies, the identification of the following information may assist the organization in determining the expected cost benefits in comparison to manual data entry:

  - color of original documents and variety of documents or form types to be identified automatically,
  - volume of hand-printed and machine-printed information to capture,
  - volume of fields per form or document,
  - volume of characters per field,
  - field type (numeric, alpha, alphanumeric),
  - extent of document preparation (pre-sorted documents, mixed form types),
  - extent of forms re-design (dropout, barcode, OMR), and
  - identification of business rules to validate or enhance the recognition result.

— **Forms management:** When the organization determines that forms management is required, it should consider both forms creation and forms processing tools. The forms creation tools should enable the authorized user to develop new forms and modify existing forms for use within a browser based application. This forms design should include the ability to create fill-in boxes, checklists, pull-down selections, free-form text input, and digital signature attachment to the form during transmission. The forms processing technologies should enable the users to manage forms using version control and support the ability to either store the submitted data with the form or store the data with the version number of the form. This information should be stored in the application database for further management and/or storage.

— **Web publishing components:** When the organization requires publishing documents to a web server, the system should support the ability for authorized users to create templates associated with specific classes or types of documents. These templates should be used by the web publishing system to convert submitted documents to either HTML or XML format including graphic and table conversion as required. The system should provide a mechanism for authorized users to either configure the system to automatically publish these converted documents directly to the web server or send the converted document to a webmaster for review and website updating.

6.8 Forms Review and Design Considerations

Most organizations utilize both electronic and hardcopy forms. These forms should be reviewed to identify where hardcopy forms could be converted to an electronic format, and electronic forms should be reviewed to identify components that can be automated along with identifying components that should stored/retrieved from legacy and ECM systems.

When reviewing electronic and hardcopy forms, the organization should consider:

- Does the form need to be an exact replica of the original hardcopy version?
- Does the form need to be printed and completed manually by outside personnel and then re-entered into the system?
- Can the form be pre-printed with information for use by outside personnel (e.g., name, address, other fields completed from information already on file, etc.)?
- Can the form include barcodes or other "coded" information that would support automated indexing after return of the completed document by outside personnel?
- Can the fields be streamlined and have components removed (e.g., received by, route to, date received, etc.)?
- Can the form be provided electronically or does it need to be in hardcopy format?

As organizations review forms, additional considerations such as whether to use forms creation applications designed primarily for web based applications or primarily for ECM environments. The use of electronic forms can greatly streamline organizational effectiveness and reduce duplicate data entry otherwise seen when using hardcopy forms and/or forms that are completed manually and then scanned/indexed.

Organizations should consider the process that will be required to create the form, deliver the form to the user, and how to accept the form and/or information after being completed by the user(s).

6.9 Legacy Data/Document Conversion Methodology Considerations

Legacy data is commonly loaded into new EDMS systems including both documents to be converted into a digital format and digitally born data. Issues associated with digitally born data loading include what versions should be loaded and how to establish common and accurate metadata. It is important to recognize that not all digitally born data will be stored in the system or stored in the native format due to age or format/structure of the data. For example, documents stored on a network drive or mainframe storage system may be in old or proprietary formats that require specialized formats and software that is no longer available or easily accessible. These types of digital data should be converted to an industry standard format such as PDF/A, HTML, or XML (or JPG, TIFF, etc. if image/map data).

For documents in hardcopy format, there are three different approaches to existing file/data conversion in use throughout the document management/workflow industries: full backfile, partial backfile, and as-needed. The organization should review and determine which approach best meets the previously defined business and technical goals. The approach selected by the organization will become extremely important if there are existing documents/files that need to be converted along with new and ongoing document receipt. Full backfile and partial backfile conversions typically require the selection of an outside conversion organization capable of processing large volumes of documents within a short time frame. The determination of whether to use an outside conversion organization or to convert using internal resources should be based on the volume of information to be scanned, the complexity of the required indexing, and the required expediency of the conversion. The three approaches are discussed more fully below.

6.9.1 Full Backfile Conversion

When selecting a full backfile conversion, the organizational goal would be to have all existing hardcopy documents available for use within the system in an electronic format. This conversion methodology is used when existing documents must be converted to meet business and/or technical goals. This

methodology is typically very expensive and time consuming. The costs associated with full backfile conversions are based on the volume of documents being converted and the total number of "keystrokes" needed to index each document, which is calculated by the total number of characters. When calculating the total number of characters, the organization should determine the level of accuracy required. For conversions where the conversion organization will only enter the information once (minimal data verification), the accuracy is typically not high enough to directly import the information into the document-imaging part of the system. It is recommended that a verification process (commonly achieved through "double keying") be implemented, which increases the cost of conversion from an industry average of $0.10 per page to $0.20 per page.

6.9.2 Partial Backfile Conversion

This conversion methodology is similar to the full backfile conversion except that the organization selects specific documents requiring conversion such as by document age or date. Other than reducing the total number of documents requiring conversion, all considerations outlined within the full backfile methodology apply.

6.9.3 As-Needed Conversion

This conversion methodology would allow the organization to convert documents only when required to complete an activity or process when new work is initiated. This conversion effort typically does not require the utilization of an outsourcing organization. To perform this type of conversion, the system should have a common "list" of where all documents are located, including both hardcopy and electronic copies. The purpose of this list is to enable the users to quickly locate documents and determine whether they are available in the document imaging system or whether they are in hardcopy format and require conversion. As new work items are received, the system should notify the user (or scan/index operator) that other documents are in hardcopy format and need to be retrieved, scanned, and indexed, prior to routing to the user(s) for processing.

6.10 Procurement Document Preparation

The procurement document provides detailed information on all aspects of the project. It is critical to include detailed technical requirements in any procurement document including information on existing and anticipated operations, along with documenting data volumes, indexing requirements, and routing requirements, etc. The technical requirements should include all necessary information allowing solution/product suppliers and integrators to respond in sufficient detail for the organization to be able to select the best product/supplier that meets the enterprise goals of the organization.

This document should also include detailed information related to the acceptance testing criteria used to validate components and solutions implemented by the selected vendors/suppliers. At a minimum, the software procurement documentation should include information on the following topics:

- Business requirements
- Technical requirements
- Software vendor experience requirements
- Identification of standards that vendor products must comply with
- Description of how the organization will conduct acceptance testing
- Requirements of user acceptance testing documentation to be prepared by the selected solution vendor or product supplier

- Software vendor project management requirements
- Description of records management capability requirements
- Staff orientation and training requirements
- Requirements for vendor supplied technical support

6.11 Solution/Product Evaluation Guidelines

When evaluating solutions/products, the organization should consider several factors associated with the product and technology including:

- **Product maturity:** The organization should evaluate the level of product maturity. This evaluation should include determining how long the product has been generally available, specifically whether the product is in an early release stage (i.e., is this a new version which has not been fully implemented by the user community yet) or whether the selected product has been in production for at least one year. All products are continually being updated to provide new functionality, "bug" fixes, and adherence to new standards and technologies. It is important for the organization to consider the maturity of each part of the selected solution when determining the overall risk factors associated with implementing these technologies.

- **Adherence to relevant industry standard/guidelines:** When reviewing various products and technologies, the organization should consider whether the selected product(s) adhere to the appropriate standards and/or guidelines.

- **Ability to meet key objectives and critical success factors:** Each organization should evaluate whether the selected product meets all, or a part, of the previously defined critical success factors. It is important that the organization select the most appropriate solution to address the previously defined business and technical requirements, rather than being forced to modify business/technical goals to meet the capabilities of the selected product. For those areas where the selected technology does not meet the stated requirements, the organization should evaluate and determine the potential risk associated with changing the requirements. Changes to requirements may be in order due to technology not being mature, the requirement being a future item, the requirement not being critical to the success of the organization, etc.

- **Level of available technical support both during implementation and after:** When selecting the product/technology, the organization should review the level of technical support both during and after technology implementation. The organization should determine whether the primary product supplier provides all support (with the exception of third party development) related to the installed product or whether technical support is only available through a reseller or "partner."

- **Product scalability:** The evaluation of any technology component should include the consideration of the expected scalability of the solution based on its ability to meet future increases in processing volumes and expanded user base.

- **Availability of system documentation, including help facilities**.

- **System security:** Due to the enterprise scope security issues have over an organization, the organization should evaluate security features in compliance with the organization's internal policies and requirements. Often the ability of the product to leverage the security features of the native operating system provides a measure of protection that will alleviate concerns over proprietary implementations. The organization may also want to evaluate and weigh product

features that support managed network services over use of applications using open sessions or "captured sessions" that provide limited security.

- **System availability:** Although system availability issues are often overlooked, many government agencies now expect a defined level of availability for the entire solution. The organization should identify particular features of the product that directly contribute to system availability and identify those single points of failure in the solution that can cause a complete outage. This evaluation should be performed within the context of the risks associated with not having the solution available during normal business hours.

- **Cost of ownership:** To determine the cost of ownership of a given solution, the organization should also consider features that address basic system administrative tasks including configuration management, software distribution, addition of new users, auditing, error reporting, disaster recovery and restoration, performance measurement utilities, and management reports. Determine whether the product requires additional software and/or hardware to maintain a test, training, and development environment.

- **Reference site benchmarks:** When available, performance benchmarks from a known reference site of similar size are invaluable in determining the solution's ability to meet the expected volume of work. Evaluate the product based on its ability to meet the peak processing loads from the reference site.

Along with these items, it is highly recommended that the review team consider how well the proposed solution/product meets the business and technical requirements along with whether the solution will adequately address the business objectives and functional needs and expectations.

It is common for organizations to include members of all aspects of the organization throughout the review process of the procurement documents including reviewing vendor/integrator responses. This should include members of the records management team, business management team, legal team, and the technical support organization.

6.12 Project Planning and Execution

After the vendor/integrator has been selected, the project planning typically occurs with the establishment of dates associated with detailed technical design, user interface definition/customization, report definition, etc. This phase of the project can take a significant amount of time, but should not be rushed. There are various approaches used in the industry to address the issue that unless the users have experience working with these technologies, there are aspects and functions that will be difficult to envision.

Taking this into account, the project plan should consider what type of design, development, and review process will be implemented ensuring that the resultant system meets the user needs while recognizing the need to have end user input to ensure successful change management. The first approach is for the integrator/vendor to create an initial design, or prototype, and have the users "test" the system to gain a better understanding of how they can use the selected technologies. The process of updating the prototype and user testing continues in this project model until the system fully meets the needs of the organization. While this ensures a system that fully meets the needs of the users during initial rollout, there are several problems associated with this to be considered including the potential for project/cost over-runs and project timelines being greatly or significantly altered and/or expanded from original expectations. The second approach is for the integrator/vendor to prepare a phased design and rollout plan that encompasses both prototype design and user review/feedback during each project phase or component to be designed/implemented.

The organization should decide the best vehicle for having the organizational members of the design team learn how the selected technologies work and the best approach to ensure user acceptance of the new technology(ies), while keeping within project budget and timeframes.

6.13 System, Unit Testing, and Project Monitoring

During the development and configuration phase of the project, it is imperative that the system be fully and properly tested. It is critical to the success of any EDMS project that full testing is performed by the vendor/integrator prior to end user review/participation. Without this thorough testing, the potential for user frustration with the system being non-functional or "not ready" is significant and can have a negative impact on other change management activities and expectations.

This testing and project monitoring should be executed by different members of the project team with participation as recommended in Table 1.

Table 1: System Testing Tasks and Roles

Integrator/Vendor Testing	Organizational Technical Testing	Organizational Business Unit Testing
Unit/Module testing – As each unit/module is developed, the integrator/vendor should fully test all aspects.	Updated on progress of unit/module testing to monitor project progress.	Updated on issues and design considerations to ensure system meets end user needs and requirements.
System Testing – As the system is configured on the test platform, the integrator/vendor should fully test all aspects of the system.	Updated on progress of system testing to monitor project progress.	Updated on project progress while continuing to prepare for implementation and addressing organizational change related issues.
	Full System Testing – After the integrator/vendor has successfully completed their testing, the organizational technical team should perform detailed system testing to ensure compliance with organizational expectations and functional/technical requirements.	
		End User Testing – After the organizational technical team has successfully completed all system testing, selected end users should test the system from the business perspective to ensure all functions operate as anticipated.

6.14 Acceptance Testing Criteria

It is recommended that the associated tests to be used to validate the system should be based on the concept that a team representing all parties would be formed. This team, including both product suppliers, end users, and project management, should be present and work together throughout the various phases

of the testing. There are other methods that can be considered to perform acceptance testing including the product supplier developing and performing the acceptance tests or the users developing and performing the tests.

To ensure that each part of this system is properly tested and that all parts of the system being implemented meet or exceed system designs (with agreed upon modifications), both the organization and the implementation team should participate in the acceptance testing and sign-off's. Those components being validated and verified include:

- verifying all system functionality is operational,
- verifying system backup and recovery procedures operate properly, and
- verifying system design specifications are met including agreed upon modifications.

This testing used should ensure that:

- the implemented system either meets or exceeds the system design documentation, and
- all users can access and use the system.

Listed below are the guidelines that should be used during the system and user testing time periods.

1. The organization should maintain a journal of events for the duration of the acceptance test and identify any hardware/software deficiencies to the product supplier.
2. No hardware or software modifications should be allowed without the approval of the project director(s)/sponsor and/or project manager. The organization should provide a reasonable but limited amount of time for overcoming problems encountered during the acceptance test.
3. Suspension of the acceptance test should occur only by mutual agreement or if the organization determines that the solution is not ready for testing. If this should occur, a re-test date should be scheduled when the product supplier is able to update the necessary components identified to be deficient.
4. At the end of the acceptance test, the project manager should review the list of deficiencies, if any, and make a determination to:
    a. accept the system based on the acceptance test results with the deficiency list, in which case the items on the list must be corrected by a mutually agreed upon date; or
    b. reject the system based on the acceptance test results, in which case the items on the deficiency list must be corrected prior to a re-test, and another site acceptance test scheduled.

6.15 Rollout Planning

When the organization completes the acceptance testing, the planning of the technology rollout should include evaluating current and planned organizational activities including other projects, ongoing work activities, and the change management issues that can affect the overall implementation. The organization should consider whether to integrate the system into a production mode using a phased approach following a "process" model or a "unit" model. The "process" model incorporates rolling out the application to all users associated with either a specific activity or group of activities. The "unit" model incorporates rolling out the application on a complete unit basis. If the organization is implementing either

document imaging or document services, the rollout plan should be based on a unit basis. When the organization is implementing workflow technologies, it should consider rolling out the application following the process model to ensure that all users have access to the electronic information. If the organization implements the workflow technology on a unit basis, caution should be exercised to ensure that users who are not within the selected organization/unit will have access to the hardcopy documents to continue/finish the work process. This is important as once the organization begins managing and processing work in an electronic environment, the hardcopy documents (previously scanned) would not be readily available.

An important aspect of rollout planning is related to user training. Organizations should ensure that sufficient and detailed end user and administrative training have been provided prior to system rollout. This training should enable the users to fully utilize the system after rollout.

6.16 Business Practices Documentation

Prior to the system being moved into full production, it is highly recommended that the organization prepare a business practices or policy document. This document further enables the organization to authenticate, or certify, that information contained within the digital system is accurate, reliable, and trustworthy. Information which should be contained in this document includes, but is not limited to, a:

- description of how information will be scanned, indexed, and verified;
- description of how the system will be secured from unauthorized access;
- description of how documents will be secured from unauthorized modification or alternation;
- description of how authorized modification of documents will be managed, including audit trail information and the ability to retrieve any previous document version required to be maintained;
- description of how notes and annotations (if any) will be stored and managed, if they are a part of the business record;
- description of how these policies and procedures will be followed; and
- description of how the system will adhere to the published records retention schedule.

All personnel using the system should follow this business practices/policy document. As changes to the system are implemented, this document should also be updated to reflect system modifications. Changes to this document should be clearly marked to denote when the change took effect and what areas were affected.

6.17 ECM Project Activities Best Practices

Detailed in Annex B, *Summary Listing of Industry Best Practices Associated with ECM Project Phases/Activities,* are industry standard activities and project steps that should be followed by the organization when developing the project plan and schedule. This list of activities should be customized as appropriate to meet organizational requirements and procurement procedures.

Annex A
(informative)
Guidelines and Standards

A.1 General

This section of the document provides detailed information on those guidelines and standards that are recommended for use when implementing EDMS. As these guidelines and standards are reviewed, the user should determine which guideline(s) and/or standard(s) would be beneficial to the organization. Copies of all referenced guidelines and standards are available through AIIM.

These guidelines and standards have been organized into five sections:

- Document Management Industry Guidelines
- Document Services Industry Standards
- Workflow Industry Standards
- Document Imaging Industry Standards
- Storage and Archival Standards

A.2 Document Management Industry Guidelines

ISO/DTS 12032, *Document management – Statistical sampling for document images*

ISO/TS 12033:2001, *Electronic imaging – Guidance for selection of document image compression methods*

ISO/TR 12037:1998, *Electronic imaging – Recommendations for the expungement of information recorded on write-once optical media*

ISO 12651:1999, *Electronic imaging – Vocabulary*

ISO 12653-1:2000, *Electronic imaging – Test target for the black-and-white scanning of office documents – Part 1: Characteristics*

ISO 12653-2:2000, *Electronic imaging – Test target for the black-and-white scanning of office documents – Part 2: Method of use*

ISO/TR 12654:1997, *Electronic imaging – Recommendations for the management of electronic recording systems for the recording of documents that may be required as evidence, on WORM optical disk*

ISO/TR 14105:2001, *Electronic imaging – Human and organizational issues for successful Electronic Image Management (EIM) implementation*

ISO/TR 15489-2, *Information and documentation – Records management – Part 2: Guidelines*

ISO/TR 15801:2004, *Electronic imaging – Information stored electronically – Recommendations for trustworthiness and reliability*

ISO 15801:2004 describes the implementation and operation of information management systems which store information electronically where the issues of trustworthiness, reliability, authenticity, and integrity are important. The whole life cycle of a stored electronic document is covered, from initial capture to eventual destruction.

This document is for use with any information management system, including traditional document imaging, workflow, and COLD/ERM technologies, and using any type of electronic storage medium including WORM and rewritable technologies.

ISO 15801:2004 does not cover processes used to evaluate the authenticity of information prior to it being stored or imported into the system. However, it can be used to demonstrate that output from the system is a true reproduction of the original document.

ANSI/AIIM TR2-1998, *Glossary of Document Technologies*

This glossary has been prepared to standardize the use of and meaning of terms associated with micrographics, electronic imaging, workflow, and related telecommunications/Internet technologies and to provide an accurate, understandable guide for both the beginner and expert. The total number of terms included has been substantially increased, although many obsolete terms from the previous edition have been eliminated. In addition, the definitions for the terms retained in this edition have been reviewed and revised as necessary to more clearly reflect present-day terminology.

ANSI/AIIM TR15-1997, *Planning Considerations, Addressing Preparation of Documents for Image Capture*

The purpose of this technical report is to provide information to organizations considering image capture as a means of converting an existing records collection. This technical report identifies possible issues that can be encountered when preparing documents for image capture. Moreover, the purpose of this report is to provide the insight necessary for quality document preparation.

ANSI/AIIM TR21-1991, *Recommendations for the Identifying Information to be Placed on Write-Once-Read-Many (WORM) and Rewritable Optical Disk (OD) Cartridge Label(s) and Optical Disk Cartridge Packaging (Shipping Containers)*

This technical report outlines recommended information that should be placed on optical disk cartridges and optical disk cartridge packaging (on a physical label or other printed surface) for the purpose of identifying the optical disk. It applies to all sizes of optical disk cartridges that can store user-recordable information. This technical report does not attempt to specify the types of container(s) or protection needed for packaging optical disks. This report is meant to give guidance to the manufacturer, supplier, and user by providing labeling and identification related recommendations.

ANSI/AIIM TR25-1995, *The Use of Optical Disks for Public Records*

This technical report was funded by a grant from the National Historic Records and Publications Commission. It is intended for federal, state, and local government agencies and related entities with records management responsibilities. In recent years, a number of government agencies have considered using electronic document imaging systems and optical disk technology for records management applications. This report provides guidelines for the planning, implementation, and operation of such systems in applications involving long-term and permanent public records.

ANSI/AIIM TR27-1996, *Electronic Imaging Request for Proposal (RFP) Guidelines*

This technical report provides guidelines for developing Request for Proposals (RFP's) for Electronic Image Management (EIM) systems that are used for document storage and retrieval and for systems used for document storage and retrieval in non-EIM environments, i.e., non-digital imaging applications. These guidelines provide step-by-step procedures for analyzing system requirements, developing functional specifications, and evaluating configuration alternatives. Guidelines have also been included for developing the administrative sections of an RFP. Office-type documents are the primary focus of this technical report. The specialized needs for engineering drawings and other document types are not considered. However, the basic principles for developing an RFP that are outlined in this document apply to a variety of electronic image-based projects. This report is being updated.

ANSI/AIIM TR28-1991, *The Expungement of Information Recorded on Optical Write-Once-Read-Many (WORM) Systems*

This technical report applies to the removal of information recorded on WORM disk media when expungement orders are ordered by the court or administrative authority; expungement requires the elimination of information. This report establishes uniform practices for both information removal and to document the action for removal. Following these recommendations will ensure that the expungement is performed consistently. This technical report does not address CD-ROM or rewritable optical media or information that is retained, managed, or distributed to satisfy the Freedom of Information Act or Privacy Act objectives.

ANSI/AIIM TR31-2004, *Legal Acceptance of Records Produced by Information Technology Systems as Evidence*

This technical report addresses laws that affect personal or business recordkeeping practices. In particular, it addresses laws containing recordkeeping provisions that require records to be kept available for government audit, require records to be submitted to the government, or establish the form of records. Includes three parts: 1) Evidence law and guidelines for performance; 2) Performance guidelines for the legal acceptance of records produces by information technology systems, and 3) Implementation – Self assessment.

ANSI/AIIM TR32-1994, *Paper Forms Design Optimization for Electronic Image Management (EIM)*

The purpose of this technical report is to provide information on characteristics of printed forms that will make them readily accepted in various EIM applications. This document covers forms characteristics that affect scanning. It also addresses forms layout, recognition technology, scanner performance, and data processing and the effect on data capture and data storage. This technical report is not intended to address forms removal technologies or the design of electronic forms.

ANSI/AIIM TR33-1998, *Selecting an Appropriate Image Compression Method to Match User Requirements*

The purpose of this technical report is to provide practical methods for analyzing user requirements for image compression in order to select an appropriate and optimal image compression scheme which matches user requirements. For example, an EIM system configured for scanning, storing, and delivering halftone, line art, text, and continuous tone images will have different image compression requirements as compared to an application involving only text. This technical report is designed to provide guidance in selecting applicable compression algorithms for each among a wide range of source documents.

ANSI/AIIM TR34-1996, *Sampling Procedures for Inspection by Attributes of Images in Electronic Image Management (EIM) and Micrographics Systems*

This technical report contains procedures that may be used to select and apply sampling inspection plans to determine if a lot or batch of electronic or micrographic images meets specified quality requirements. Its purpose is to do the following:

— provide guidance to the user when selecting a sampling procedure that will meet risk requirements, and

— enable the user to develop a sampling plan for individual images in a scientific manner.

ANSI/AIIM TR35-1995, *Human and Organizational Issues for Successful EIM System Implementation*

This document provides a fundamental framework for understanding the basic issues and concepts of organizational factors, human factors, and ergonomics for Electronic Image Management (EIM) systems. The principles of human factors and ergonomics are applied to usability criteria for the development and selection of EIM equipment, environmental and implementation issues, and training for long-term productivity benefits. This technical report should help readers understand and plan for the non-technical issues that need to be managed when implementing EIM. Recommendations are provided to help prepare organizations for change.

ANSI/AIIM TR40-1995, *Suggested Index Fields for Documents in Electronic Image (EIM) Environments*

The purpose of this technical report is to describe fields of attribute information that are often used with electronic imaging systems. This information may take the form of a collection of database fields or a structured computer record that refers to an image record on an electronic, digital image medium. Such a collection of database fields includes a necessary and sufficient description of the image record to control subsequent storage, retrieval, and archive management related actions with that image record. The information contained in the fields described in this document is similar to that typically used in a text management system. It is designed to be informative to a user if it is displayed in an image query response. System designers could elect to use some or all of the fields described in this technical report in addition to fields that are specific to the application they are designing.

ANSI/AIIM TR41-2006, *Optical Disk Storage Technology, Management, and Standards*

This technical report provides information on the various technologies, management, implementation strategies, and industry standards for optical based subsystems. This information and the corresponding techniques described have been provided to enable optical disk system users, as well as other imaging system users, to become knowledgeable in the various techniques currently in use throughout the imaging industry

A.3 Document Services Industry Standards

When reviewing document services technologies you should determine whether or not these products meet the recommended industry standards. A vendor/supplier will be able to tell you if they are certified for the following industry standards.

A.3.1 Document Management Alliance (DMA) 1.0 Specification

The DMA specification defines software component interfaces that enable uniform search and access to documents stored in multi-vendor document management systems. The DMA organization included more than 60 user and vendor companies working together as a task force to define interoperability specifications that meet the requirements of enterprise document management systems.

A.3.2 DMWare

DMWare is the open-source distribution and development clearinghouse for document management. The subject matter of DMWare, based on the work of the Document Management Alliance (DMA) and of the Open Document Management API (ODMA) coalition, is public, openly contributed document management software, documentation, and metadata definitions.

A.3.3 Open Document Management API (ODMA)

ODMA specifies a set of interfaces that applications can use to initiate actions within a document management system. The API is intended to be relatively easy for application vendors to incorporate into updates of existing applications. It should not require major restructuring of an application to integrate it with ODMA.

NOTE This version of ODMA does not specify how document management systems may initiate actions within the applications.

A.4 Workflow Industry Standards

WfMC-TC-1009, Interface 2 Workflow Client Application Programming Interface (Interface 2 & 3) Specification

The purpose of this document is to specify standard workflow management Application Programming Interfaces (APIs) which can be supported by workflow management (WFM) products. These API calls provide for a consistent method of access to WFM function in cross-product WFM engines. The API set is named Workflow Application Programming Interfaces (WAPI).

WfMC-TC-1015, Audit Data Specification

The purpose of this document is to specify what information needs to be captured and recorded from the various events occurring during a workflow enactment. This document does not define how the data is stored, but what information is to be gathered and made available for analysis. The information will be called Common Workflow Audit Data (CWAD).

WfMC-TC-1018, Interoperability, Internet, e-mail MIME Binding

This document maps to the WfMC standard, Interoperability Abstract Specification (WfMC-TC-1012), which provides an abstract specification that defines the functionality necessary to achieve a defined level of interoperability between two or more workflow engines. This document defines a binding that gives concrete type definitions and message formats for the realizations of the abstract specification, using Internet e-mail with MIME encoding as the transport mechanism.

A.5 Document Imaging Industry Standards

ISO 10196:2003, *Document imaging applications – Recommendations for the creation of original documents*

ISO 12653-1:2000, *Electronic imaging – Test target for the black-and-white scanning of office documents – Part 1: Characteristics*

ISO 12653-2:2000, *Electronic imaging – Test target for the black-and-white scanning of office documents – Part 2: Method of use*

ANSI/AIIM MS44-1993, *Recommended Practice for Quality Control of Image Scanners*

This recommended practice provides procedures for the ongoing control of quality within a digital document image management system. The objective is to provide a means of quality control from input to output. Regular use of the recommended procedures should ensure continued maintenance of an established level of quality.

ANSI/AIIM MS52-1991, *Recommended Practice for the Requirements and Characteristics of Original Documents Intended for Optical Scanning*

This standard describes the physical characteristics of paper documents which facilitate black-and-white optical scanning and the characteristics which make scanning either difficult or impossible. It provides general recommendations for the design of documents in order to make these documents easier to scan. This standard does not cover specific scanning applications, such as scanning of checks, scanning of engineering drawings, or scanning of barcodes, which are the subject of other standards. It does not address the technical details for OCR, which are the subject of other standards. Moreover, oversized documents and tiling techniques are not specifically addressed in this standard, although many of the same principles apply.

ANSI/AIIM MS53-1993, *Recommended Practice – File Format for Storage and Exchange of Image; Bi-Level Image File Format: Part 1*

The purpose of this standard is to standardize a self-contained file format for the transfer of bi-level image files in environments other than facsimile telecommunications. The image file format is similar to a Document Application Profile (DAP) and supports the transfer of encoded bi-level raster scan images in environments. This standard covers bi-level images that are coded using CCITT T.4 (Group 3) and T.6 (Group 4), as well as bit-mapped images (having no compression). The file format is media independent.

ANSI/AIIM MS55-1994, *Recommended Practice for the Identification and Indexing of Page Components (Zones) for Automated Processing in an Electronic Image Management (EIM) Environment*

This document identifies a media and application independent structure and indexing scheme that will allow necessary and sufficient description of document pages and zones (rectangular sub areas) within a page. These zones can then be processed automatically in the most appropriate fashion, regardless of the nature of data outside the identified zone(s). In particular, this standard recommended practice defines a document page so that the following processes can be applied to its electronic image record:

- data compression specifically suitable to the nature of the data within the zone (e.g., JPEG compression, vs. T.6 compression used in Group 4 Fax);
- optical mark recognition;
- optical character recognition;
- intelligent character recognition;
- handprint character recognition;
- raster-to-vector conversion for computer aided design (CAD) or geographic information system (GIS) applications;
- signature capture and recognition (CSR); and
- any other form of compression, image manipulation or pattern recognition technology, or algorithm(s) that may rely on specific data capture or storage methods.

A.6 Storage and Archival Standards

This Storage and Archival section is divided into several storage technologies. The storage technologies included are:

- Technical Reports (general)
- Magnetic WORM
- Blue Laser Optical Storage
- Red Laser Optical Storage

A.6.1 Storage and Archival Technical Reports (General)

ANSI/AIIM TR41-2006, *Optical Disk Storage Technology, Management, and Standards*

This technical report provides information on the various technologies, management, implementation strategies, and industry standards for optical based subsystems. This information and the corresponding techniques described have been provided to enable optical disk system users, as well as other imaging system users, to become knowledgeable in the various techniques currently in use throughout the imaging industry

A.6.2 Magnetic WORM Storage and Archival Standards

As of 2005, no national (ANSI, AIIM) or international (ISO) storage or archival standards have been published.

A.6.3 Blue Laser Optical Storage and Archival Standards

ISO/IEC 17345:2006, ***Information technology – Data interchange on 130 mm Rewritable and Write Once Read Many Ultra Density Optical (UDO) Disk Cartridges – Capacity: 30 Gbytes per Cartridge – First Generation***

A.6.4 Red Laser Optical Storage and Archival Standards

ANSI/INCITS/ISO/IEC 10089:1991 (R2004), ***Information technology – 130 mm Rewritable Optical Disk Cartridge for Information Interchange***

ISO/IEC 13549:1993, ***Information technology – Data interchange on 130 mm optical disk cartridges — Capacity: 1,3 gigabytes per cartridge***

ISO/IEC 11560:1992, ***Information technology – Information interchange on 130 mm optical disk cartridges using the magneto-optical effect, for write once, read multiple functionality***

ISO/IEC 14517:1996, ***Information technology – 130 mm optical disk cartridges for information interchange — Capacity: 2,6 Gbytes per cartridge***

ISO/IEC 15286:1999, ***Information technology – 130 mm optical disk cartridges for information interchange — Capacity: 5,2 Gbytes per cartridge***

ANSI/INCITS 220-1992 (R2002), ***130-mm Optical Disk Cartridges of the Write-Once, Read Multiple (WORM) Type, Using the Magnetic-Optical Effect for Write Once, Read Multiple Functionality***

ANSI/INCITS 234-1993 (R2005), ***Test Methods for Media Characteristics of 130 mm Re-writable Optical Digital Data Disks with Continuous Composite Servo (CCS)***

Annex B
(informative)
Summary Listing of Industry Best Practices Associated with ECM Project Phases/Activities

Activity	Start Date	End Date
Process/Procedure Base Lining		
Anticipated Processes/Procedures		
Technology Requirement Definition		
Document Classification and Indexing		
Business Objectives and Requirements		
Technology Evaluation		
Forms Evaluation and Re-design (if needed)		
Legacy Data/Document Conversion Methodology Considerations		
Digitally Born Data Import/Management Requirements		
Storage Technology Requirements for Short/Medium/Long Term Storage		
User Access/Functionality Requirements		
Acceptance Testing Criteria from the Technical and Business Perspectives		
Procurement Documents (RFP, etc.)		
Vendor/Product Evaluation & Selection (including reviewing other similar environments where the proposed solution/system has been implemented)		
Detail Application Design (including legacy system integration)		
System Development/Implementation		
Business/Policy/Procedure Documentation		
Unit and System Testing and Project Monitoring		
User Training (for users participating in initial testing and general user training)		
Acceptance Testing (from both the business and technical perspectives)		
Phase System Rollout (as appropriate for the organization, recognizing the time required and necessary to properly manage the change in the way the organization conducts business and manages digital information)		

AIIM/ARP 1-2007

1100 Wayne Avenue, Suite 1100
Silver Spring, MD 20910-5603
Tel: 301/587-8202
FAX: 301/587-8202
e-mail: aiim@aiim.org
Website: http://standards.aiim.org
ISBN 0-89258-415-7
Printed in the United States of America